DOCUMENT RESUME

ED 117 130 TM 004 982

AUTHOR Tritchler, D. L.; Pedrini, D. T.
TITLE CORCOR: Discussion and Listing of a Computer Program to Correct the Correlation Coefficient for Restriction of Range.
PUB DATE Sep 75
NOTE 9p.

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS *Computer Programs; *Correlation; *Sampling; Statistical Analysis
IDENTIFIERS *CORCOR

ABSTRACT

When a correlation is taken in a sample in which the range of the variables being correlated is limited, it is often desired to apply the correlation to groups with different ranges for the variables. In such situations, a correction for restriction of range (more precisely, restriction of standard deviation) may be necessary. Considered in this paper are four cases in which correction on the correlation coefficient is possible and a computer program to perform the correlations. The program, which includes an interactive correction algorithm for explicit restriction on both variables, is listed in the article. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117130

September 1975

CORCOR:
DISCUSSION AND LISTING OF
A COMPUTER PROGRAM TO
CORRECT THE CORRELATION COEFFICIENT
FOR RESTRICTION OF RANGE

D.L. TRITCHLER AND D.T. PEDRINI
University of Nebraska at Omaha

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Abstract

When a correlation is taken in a sample in which the range of the variables being correlated is limited, it is often desired to apply the correlation to groups with different ranges for the variables. In such situations a correction for restriction of range (more precisely, restriction of standard deviation) may be necessary. Considered in this paper are 4 cases in which correction on the correlation coefficient is possible and a computer program to perform the correlations. The program, which includes an interative correction algorithm for explicit restriction on both variables, is listed in the article.

TM004 982

INTRODUCTION

The Pearson product - moment correlation coefficient (r) in a restricted group is almost invariably smaller than it would be in an unrestricted group, if the restriction was by non-random selection on the basis of one or both of the variables being correlated, or a related third variable. This often occurs when a group is determined by dropping individuals below a specified cutoff score on one of the variables. This article presents a computer program which corrects the Pearson r for 4 cases of restriction of standard deviation.

Statistical Discussion

Selecting a group on the basis of their scores on a predictor or a criterion usually reduces the range of scores to a fraction of the possible range of scores, reducing the standard deviations (SDs). The Pearson r, which is directly related to the magnitude of the SDs, is consequently reduced. If the unrestricted SDs are known, in some cases it is possible to estimate r for the unrestricted group, given the obtained r and SDs. Figure 1 shows 4 cases which have correction procedures for estimating the unrestricted r. Guilford (1965) discusses and gives formulas for cases I, II and III. Corrections for cases I, II and IV are given by Wells and Fruchter (1970).

CASE	RESTRICTION PRODUCED BY	S KNOWN FOR	s KNOWN FOR	r KNOWN FOR
I	Selection on basis of X1	X1	X1	R12
II	Selection on basis of X1	X2	X2	R12
III	Selection on basis of X3	X3	X3	R12, R13, R23
IV	Selection on basis of X1 and X2	X1, X2	X1, X2	R12

S is the SD for the unrestricted group.
s is the SD for the restricted group.

Figure 1. Correction for R12.

The Computer Program

Computer program CORCOR will compute corrected r's for each of the 4 cases of Table 1. The program was written in Fortran IV and compiled and executed on an IBM 360/65 in batch mode. However, this application also is well suited to interactive mode.

Input

The data is entered according to Figure 2, one data card for each problem.

		CARD COLUMNS				
	1	2-11	12-21	22-31	32-41	42-51
1	1	RESTRICTED COEFFICIENT	RESTRICTED STD. DEV.	UNRESTRICTED STD. DEV.	BLANK	BLANK
2	2	R. COEF.	R. SD	U. SD	BLANK	BLANK
3	3	R. COEF.	R. SD	U. SD	R13	R23
4	4	R. COEF.	R. SD	U. SD	R. SD	U. SD

All Fields except col. 1 are FIO.4
Any number of data cards may be submitted
After all the data cards, include a card with a 9 in col. 1

Figure 2. Data format for program CORCOR.

Output

Output consists of an 'echo' of the input data and the corrected correlation coefficient. A sample output is shown in Figure 3.

```
          CORRECTION OF R FOR CASE 3

Correlation with the restriction is        0.3000
The correlations with the 3rd variable are         0.4000      0.7000
The restricted S.D. of the 3rd var. is    20.0000
The unrestricted S.D. is     30.0000

Corrected R is      0.4673
```

Figure 3. Sample output for program CORCOR.

Capabilities and Limitations

Any number of corrections may be performed on a given run. Any value of r is correctable for cases 1, 3 and 4. For case 2, the formula breaks down for r=Ø.

PROGRAM LISTING 08/15/75

```
C                                                                       COR00001
C                                                                       COR00002
C            PROGRAM WRITTEN FEB. 1973 BY DLT                           COR00003
C     THE CORRECTIONS FOR CASES 1,2,3 ARE GIVEN IN GUILFORD. CASE 4 IS  COR00004
C   GIVEN IN WELLS AND FRUCHTER(1970)                                   COR00005
C                                                                       COR00006
C                                                                       COR00007
C      THE DATA IS FORMATED ACCORDING TO                                COR00008
C     THE MATRIX GIVEN BELOW.                                           COR00009
C                                                                       COR00010
C                              CARD COLUMNS                             COR00011
C        1    2-11           12-21          22-31          32-41      42-51  COR00012
C      ************************************************************     COR00013
C     1 * 1  RESTRICTED   RESTRICTED   UNRESTRICTED     BLANK      BLANK  COR00014
C        *    COEFFICIENT  STD. DEV.    STD. DEV.                        COR00015
C  C     *                                                              COR00016
C  A  2 * 2    R. COEF.     R. SD          U. SD          BLANK      BLANK  COR00017
C  S     *                                                              COR00018
C  E     *                                                              COR00019
C     3 * 3    R. COEF.     R. SD          U. SD          R13         R23  COR00020
C        *                                                              COR00021
C        *                                                              COR00022
C     4 * 4    R. COEF.     R. SD          U. SD          R. SD      U. SD  COR00023
C                                                                       COR00024
C                                                                       COR00025
C                                                                       COR00026
C                                                                       COR00027
C      ALL FIELDS EXCEPT COL. 1 ARE F10.4                               COR00028
C      ANY NUMBER OF DATA CARDS MAY BE SUBMITTED                        COR00029
C      AFTER ALL THE DATA CARDS,INCLUDE A CARD WITH A 9 IN COL 1        COR00030
C                                                                       COR00031
C                                                                       COR00032
      COMMON R12,RORS(4)                                                COR00033
      REAL R12,RORS                                                     COR00034
      INTEGER TYPE                                                      COR00035
    6 READ(5,150,END=200) TYPE,R12,RORS                                 COR00036
  150 FORMAT(I1,5F10.4)                                                 COR00037
      IF(TYPE.EQ.9) GO TO 200                                           COR00038
      GO TO (1,2,3,4),TYPE                                              COR00039
    1 CALL TYPE1                                                        COR00040
      GO TO 5                                                           COR00041
    2 CALL TYPE2                                                        COR00042
      GO TO 5                                                           COR00043
    3 CALL TYPE3                                                        COR00044
      GO TO 5                                                           COR00045
    4 CALL TYPE4                                                        COR00046
    5 PRINT 100                                                         COR00047
  100 FORMAT(/1X,'---------------------------------------- '//)         COR00048
      GO TO 6                                                           COR00049
  200 STOP                                                              COR00050
      END                                                               COR00051
      SUBROUTINE TYPE1                                                  COR00052
      COMMON R12,SDXR,SDXP                                              COR00053
      REAL R12,SDXR,SDXP,XR12                                           COR00054
C                                                                       COR00055
C R12 IS UNCORRECTED R,SDXR IS RESTRICTED SD,SDXP IS UNRESTRICTED S.D.  COR00056
C THIS IS CASE 1 AS GIVEN IN GUILFORD                                   COR00057
C                                                                       COR00058
      PRINT 100,R12,SDXR,SDXP                                           COR00059
  100 FORMAT(9X,'CORRECTION OF R FOR CASE 1'//1X,'CORRELATION WITH THE RCOR00060
     XESTRICTION IS',1X,F10.4/1X,'S.D. FOR EXPLICITLY RESTRICTED VAR.  ICOR00061
```

```
     XS',1X,F10.4/1X,'UNRESTRICTED S.D. IS',1X,F10.4//)                    COR00062
      XR12=RX(R12,SDXP,SDXR)                                               COR00063
      PRINT 101,XR12                                                       COR00064
  101 FORMAT(1X,'CORRECTED R IS',1X,F10.4)                                 COR00065
      RETURN                                                               COR00066
      END                                                                  COR00067
      SUBROUTINE TYPE2                                                     COR00068
      COMMON R12,SDYR,SDYP                                                 COR00069
      REAL R12,SDYR,SDYP,XR12                                              COR00070
C                                                                          COR00071
C R12 IS UNCORRECTED R,SDYR IS RESTRICTED SD,SDYP IS UNRESTRICTED S.D.     COR00072
C THIS IS CASE 2 AS GIVEN IN GUILFORD                                      COR00073
C                                                                          COR00074
      PRINT 100,R12,SDYR,SDYP                                              COR00075
  100 FORMAT(9X,'CORRECTION OF R FOR CASE 2'//1X,'CORRELATION WITH THE RCOR00076
     XESTRICTION IS',1X,F10.4/1X,'S.D. FOR IMPLICITLY RESTRICTED VAR.      COR00077
     XIS',1X,F10.4/1X,'UNRESTRICTED S.D. IS',1X,F10.4//)                   COR00078
      XR12=SQRT(1.-((SDYR*SDYR)/(SDYP*SDYP))*(1.-R12*R12))                 COR00079
      XR12=(R12/ ABS(R12))*XR12                                            COR00080
      PRINT 101,XR12                                                       COR00081
  101 FORMAT(1X,'CORRECTED R IS',1X,F10.4)                                 COR00082
      RETURN                                                               COR00083
      END                                                                  COR00084
      SUBROUTINE TYPE3                                                     COR00085
      COMMON R12,SDZR,SDZP,R13,R23                                         COR00086
      REAL R12,R13,R23,SDZR,SDZP,XR12                                      COR00087
C                                                                          COR00088
C R12 IS UNCORRECTED R,R13 AND R23 ARE CORRELATIONS WITH THE EXPLICTLY     COR00089
C RESTRICTED VARIABLE,SDZR IS RESTRICTED S.D. FOR EXPLICITLY RESTRICTED    COR00090
C GROUP,SDXP IS UNRESTRICTED S.D.                                          COR00091
C THIS IS CASE 3 AS GIVEN IN GUILFORD                                      COR00092
C                                                                          COR00093
      PRINT 100,R12,R13,R23,SDZR,SDZP                                      COR00094
  100 FORMAT(9X,'CORRECTION OF R FOR CASE 3'//1X,'CORRELATION WITH THE RCOR00095
     XESTRICTION IS',1X,F10.4/1X,'THE CORRELATIONS WITH THE 3RD VARIABLECOR00096
     X ARE',1X,F10.4,1X,F10.4/1X,'THE RESTRICTED S.D. OF THE 3RD VAR. ISCOR00097
     X',1X,F10.4/1X,'THE UNRESTRICTED S.D. IS',1X,F10.4//)                 COR00098
      FAC=(SDZP*SDZP)/(SDZR*SDZR)-1.                                       COR00099
      XR12=R12+(R12* ABS((R13*R23)/R12)*FAC)                               COR00100
      XR12=XR12/SQRT((1.+R13*R13*FAC)*(1.+R23*R23*FAC))                    COR00101
      PRINT 101,XR12                                                       COR00102
  101 FORMAT(1X,'CORRECTED R IS',1X,F10.4)                                 COR00103
      RETURN                                                               COR00104
      END                                                                  COR00105
      SUBROUTINE TYPE 4                                                    COR00106
      COMMON R12,SDXR,SDXP,SDYR,SDYP                                       COR00107
C                                                                          COR00108
C     THIS ALGORITHM IS GIVEN BY WELLS AND FRUCHTER(1970)                  COR00109
C     R12 IS UNCORRECTED R,SDXR IS SD OF RESTRICTED SAMPLE OF X VARIABLECOR00110
C     ,SDXP IS UNRESTRICTED SD,ETC.                                        COR00111
C                                                                          COR00112
      PRINT 100,R12,SDXR,SDXP,SDYR,SDYP                                    COR00113
  100 FORMAT(9X,'CORRECTION OF R FOR CASE 4'//1X,'RESTRICTED CORRELATIONCOR00114
     C COEFFICIENT',1X,F10.4/1X,'RESTRICTED AND UNRESTRICTED STD. DEVIATCOR00115
     CIONS OF FIRST VAR',1X,F10.4,1X,F10.4/1X,'RESTRICTED AND UNRESTRICTCOR00116
     CED STD. DEVIATIONS OF SECOND VAR',1X,F10.4,1X,F10.4)                 COR00117
      SXE=SDXR                                                             COR00118
    1 R1=RX(R12,SDXP,SXE)                                                  COR00119
      SYI=SI(R1,R12,SDYP)                                                  COR00120
      SYE=SDYP+SDYR-SYI                                                    COR00121
      R2=RX(R12,SDYP,SYE)                                                  COR00122
      SXI=SI(R2,R12,SDXP)                                                  COR00123
```

```
      SXN=SDXP+SDXR-SXI                                                 COR00124
      IF(SXE-SXN.LT.0.00001)GO TO 2                                     COR00125
      SXE=SXN                                                           COR00126
      GO TO 1                                                           COR00127
    2 R1=RX(R12,SDXP,SXN)                                               COR00128
      RCE=RX(R1,SDYP,SYE)                                               COR00129
      PRINT 101,RCE                                                     COR00130
  101 FORMAT(1X,'CORRECTED R IS',1X,F10.4)                              COR00131
      RETURN                                                            COR00132
      END                                                               COR00133
      FUNCTION RX (R,SMAJ,SMIN)                                         COR00134
      XNUM= R*(SMAJ/SMIN)                                               COR00135
      RX=XNUM/SQRT(1.-R*R+XNUM*XNUM)                                    COR00136
      RETURN                                                            COR00137
      END                                                               COR00138
      FUNCTION SI (RMAJ,RMIN,S)                                         COR00139
      SI=S*SQRT((1.-RMAJ*RMAJ)/(1.-RMIN*RMIN))                          COR00140
      RETURN                                                            COR00141
      END                                                               COR00142
```

JCL043 -- JOB 714 -- HASP-II OUTPUT COMPLETE AT 18:57

BIBLIOGRAPHY

Guilford, J.P. Fundamental Statistics in Psychology and Education. Fourth edition. New York: McGraw-Hill, 1965.

Kaufman, A.S. Restriction of range: questions and answers. Test Service Bulletin, 1972, No. 59.

Wells, D.G. and Fruchter, B. Correcting the correlation coefficient for explicit restriction on both variables. Educational and Psychological Measurement, 1970, 30, 925-934.